A Comparative analysis molecular motion of F1-ATPases derived from different origins using Gaussian Network

Md. Abu Taher Khan[1], Md. Hafijur Rahaman[1], Mohammad SolaimanWalid Khan[1],
Shahajan Miah[1], Md. Aminul Islam[1] and Mohammad Delawer Hossain[1]

[1]*Department of Physics, Shahjalal University of Science and Technology, Bangladesh*

Abstract: The molecular motions of F1-ATPases derived from bovine heart mitochondria (MF1) and thermoalkaliphilic bacteria (TA2F1) were studied in this paper using Gaussian Network Model (GNM)[13-15]. The data obtained fromGNM calculation using a cut off distance 10 Å were compared and analyzed. MF1 was taken as the reference protein for comparison. The calculated force constants (γ) for TA2F1 and MF1 proteins were 0.054 kcal/ (mole $\cdot$ Å^2) and 0.08 kcal/(mole $\cdot$ Å^2) respectively. The larger force constant of MF1 implies that the inter residue contact potential of MF1 is stronger than that of TA2F1. The fluctuations exhibited by the different chains of $\alpha$and βsubunits of MF1 are very high compared to the TA2F1 protein. The largest fluctuations ~ 2.5Å^{2}was observed in the carboxy terminal region of the chain B of MF1 suggesting that the $\alpha_3\beta_3$cylinder of MF1 is flexible whereas the $\alpha_3\beta_3$cylinder of TA2F1 is highly rigid. The γsubunit of TA2F1 generated a fluctuation ~ 2.5Å^2. The fluctuation of εsubunit of TA2F1 tagged with the protruding portion of the γsubunit is the highest ~3.5Å^{2}compared to all other subunits of both the proteins. The theoretical and experimental B factors match well. The correlation coefficient for MF1 was 0.71 (71%) and for TA2F1 was 0.64 (64%). In the case of both proteins, the hydrophobic 'sleeve' forming the molecular bearing was found to be located in the minima region of fluctuation curves.

Keyword:F1-ATPases, Gaussian Network Model, MF1 and TA2F1 proteins

I.   INTRODUCTION

Molecular motor proteins are nanoscale natural machines found in biological cells. They convert chemical energy derived from the hydrolysis of adenosine triphosphate (ATP) into mechanical workor motion that forms the basis of cellular function. ATP molecules serve as remarkable sources of chemical energy that is used to perform the various cellular functions. Scientists of diverse background are studying the structure and function of these motors to either to mimic their action or to incorporate them into novel devices. Motor proteins are involved in ATP synthesis, organelle transport, muscle contraction, protein folding and nucleic acid translocation. They also take part in cell signaling, division and motion. The most prominent examples of ATP-driven linear motors are myosin, kinesin and dynein. Myosin is involved mainly in muscle contraction and it slides on actin filament. Kinesin moves along microtubules and is responsible for transporting cellular cargo such as organelles and signaling molecules. Dynein is another microtubule based motor protein that does many cellular functions. Examples of rotary molecular motors are bacterial flagella motor and ATP synthase.

The rotary enzyme ATP synthase found in mitochondria, bacteria and chloroplasts generates ATP using the free energy stored in transmembrane ion gradients [1-6]. This enzyme consists of two parts (F_o and F_1). The asymmetric membrane-bound F_o that contains the proton channel and the soluble F_1 that contain three catalytic sites which work cooperatively. These two motors are connected through an asymmetric rotor shaft made of the coiled-coil gamma subunit. The proton flow through $F_o$leads to the ATP synthesis in F_1 and ATP hydrolysis in F_1 leads to the rotation of the rotor in opposite direction that finally pumps the protons back towards F_o. The F_1 portion consists of $\alpha_3\beta_3\gamma\varepsilon\delta$ subunits. F_1 only hydrolyzes ATP and is called F_1-ATPase. Both the three α and three βsubunits can bind ATP molecules but only three β subunits contain catalytic sites in which ATP can be hydrolyzed [7]. The ε subunit acts as an inhibitor of ATP hydrolysis in F_1-ATPases [8]. The δ subunit is a part of stator in ATP synthase [9]. ATP hydrolysis driven rotation of γ subunit around $\alpha_3\beta_3$ cylinder was observed in thermophilic F_1-ATPase [10]. ATP was generated in β subunits of F_1-ATPase by mechanical rotation of the γ subunit in a direction opposite to the direction of ATP hydrolysis [11].

The proteins are flexible and can adopt different conformations in response to their environmental conditions. The protein function is a dynamic property [12]. To elucidate protein function it is necessary to understanding structure and dynamics of proteins. The structures of different proteins in the folded or native state seem to be in static form but proteins are engaged in functional motions during various cellular activities. Such motion in proteins concertedly occurs within amino acids, subunits and domains. A protein in motion frequently changes its conformation to do a particular function. Many biological processes like catalysis, transport of metabolites, folding, binding, molecular recognition etc. are performed by different proteins in the presence of motion.

Highly efficient as well as physically realistic computational methods and tools are essential to understand the fluctuation dynamics of proteins. The systematic computational characterization of protein dynamics by molecular dynamics simulations is highly time consuming. Moreover, in the case of all atom models, computation is very expensive due to the use of complicated potential energy functions. In order to overcome such problems, a very simplified Gaussian Network Model (GNM) was developed [13-15] to predict the fluctuations of protein structures in coarse-grained level. This model involves normal mode analysis (NMA) to assess collective motions of proteins. The GNM was constructed based on inter-

residue contact topology in the folded state with a single parameter potential. The observed motion in a protein is dissected into a collection of modes which give information on the molecular mechanisms relevant to biological function [16, 17].

There are mainly two types of modes in the case of protein motion: slow mode and fast mode. In a slow (low frequency) mode, the different parts of a protein undergo motions with large-amplitudes. The slow mode motions are also called 'global motions' that are essential for understanding the protein function [18-19]. The slow modes make major contributions to thermal conformational fluctuations [20, 21].Such motions can influence the interactions of proteins with other molecules and its environment. The fast modes with higher frequency fluctuations are more localized that include only a few residues and can play an important role in signal transmission, enzyme reactions and other internal processes [18]. In some studies [13, 22], it was revealed that the minima in the global mode shapes coincide with the hinge sites of the protein molecule and the maxima correspond to substrate recognition sites.

Different experimental methods are used to determine the structures of proteins such as X-ray diffraction, nuclear magnetic resonance (NMR), electron microscopy etc. The X-ray crystallographic experiments give useful structural information about thermal and other fluctuations of the atoms in a protein. In X-ray crystallographic method, the average position of each atom is determined based on the electron density map and every atom is assigned to have a B-factor whose magnitude is directly proportional to the mean square displacement from its mean position. The B-factors reflect the flexibility and dynamics of protein structures. Models for protein dynamics provide insights into the mechanisms of motion of large proteins [12]. The B-factors measured by experiments can give limited information about molecular collective motions relevant to biological functions [23].

The aim of the present work is to study the fluctuation dynamics of thermoalkaliphilic F_1-ATPase (denoted as $TA2F_1$) [24] by the GNM and compare the results with the data obtained for the reference protein mitochondrial F_1-ATPase (denoted as MF_1) [7]. The structure files for MF_1 (PDB ID: 1bmf) and for $TA2F_1$ (PDB ID: 2QE7) proteins were downloaded from Protein Data Bank (PDB) [25]. The mean square equilibrium fluctuations as well as B factors were calculated using oGNM server maintained by Yang *et al.* [26, thanks to Prof. I. Bahar, University of Pittsburgh, USA for her kind permission to use the server]. The calculated B factors and fluctuations for both proteins were analyzed and compared. It was found that the calculated values of B factors agree well with the experimental ones.

The plan of this paper is given as follows: Section 2 deals with the Comparative Analysis of Data and section 3 is presented to discuss the work. Section 4 is devoted to the conclusion of the work.

II.   DATA ANALYSIS

The mean square fluctuations and B factors of thermoalkaliphilic F_1-ATPase ($TA2F_1$) and mitochondrial F_1-ATPase (MF_1) were calculated using oGNM server [26]. The data obtained by calculation was thoroughly analyzed.

2.1 Analysis of mean square fluctuations

The mean square fluctuations exhibited by the different subunits of both the MF_1 and $TA2F_1$ proteins were compared and analyzed in the following sections.

2.1.1 Analysis of the mean square fluctuations for different chains of the α subunit

Proteins take part in various cellular processes to keep the cells alive. The functions of the proteins such as catalytic activity, folding, binding and molecular recognition involve protein motions or fluctuations [8]. Determination and understanding the three dimensional structures of proteins in the folded forms is necessary because their function depend on their folded native forms. The motions of proteins or conformational changes relevant to biological events are called functional motions or functional dynamics [12]. The fluctuations involved in biological functions of proteins are generated in collective manner that consists of correlated/anti-correlated motions of large groups of atoms, residues or domains of the proteins. Studying fluctuation dynamics of proteins is thus important to know their functional mechanisms.

The Gaussian Network Model (GNM) [15, 26] is used in this thesis to theoretically calculate the mean square fluctuations and B factors of the residues of two proteins: thermoalkaliphilic F_1-ATPase ($TA2F_1$) and mitochondrial F_1-ATPase (MF_1). The GNM dissects protein motions into a collection of various normal modes and the slowest modes usually give information on the collective motions relevant to a biological function [12]. The data obtained for these two proteins from GNM calculation [26] for the 1st slowest mode of motion were analyzed. A cutoff distance $r_c = 10$ Å for inter residue interactions was used for both the proteins. It was found in a previous study that GNM results are insensitive to the cutoff distance provided in the range 7.3 Å$\leq r_C \leq 15$ Å [12]. By GNM calculation, a force constant of $\gamma = 0.054$ kcal/(mole·Å^2) was obtained for $TA2F_1$ and $\gamma = 0.08$ kcal/(mole·Å^2) was for MF_1. The force constant is a measure of the strength of intramolecular potentials that stabilize the native fold [7]. A force constant of 1.0 ± 0.5 kcal/(mol·Å^2) has been obtained in some proteins in a previous investigation [29].

The structure of thermoalkaliphilic F_1-ATPase ($TA2F_1$) consists of eight chains (A, B, C, D, E, F, G and H) and mitochondrial F_1-ATPase has seven chains (A, B, C, D, E, F and G) as discussed inthestructure of F_1-ATPase. The motion of the amino acid residues contained in different subunits of the two proteins is revealed in the fluctuation curves. In order to compare, the mean square fluctuations of the amino acid residues involved in three α subunits (A, B and C chains), three β subunits (D, E and F chains) and one γ subunit (G chain) of $TA2F_1$ and MF_1 proteins were plotted against the relevant residue indices as shown in Figs. 2.1-2.3. All these curves were plotted on the same scale of the mean square fluctuations and residue indices. The motion of the amino acid residues contained in the ε subunit (H chain) of $TA2F_1$ protein is shown in Fig. 2.4.

The fluctuations of the three α subunits (A, B and C chains) of both the $TA2F_1$ and MF_1 proteins are shown in Fig.2.1 in which there are three curves for each chain. In the top figure obtained for the chain A, the red curve represents the

fluctuations for TA2F$_1$ protein. This curve is almost straight and parallel to the residue index axis. The black line was obtained for MF$_1$ protein and it contains some significant fluctuations at different residue index positions. Comparing these two curves (red and black) on the same scale, it is clear that the fluctuation for the TA2F$_1$ protein is almost 0.5Å^2 less than that of the TA2F$_1$ protein around the carboxy terminal region. The amino terminal of TA2F$_1$ protein starts at 27 (GLU) and ends at 500 (PRO) in the carboxy terminal region and for the MF$_1$ protein, the chain starts at 24 (ASP) and ends at 510 (ALA). The green curve is an expanded version of the red curve for TA2F$_1$ protein. The purpose of expansion is to observe whether there is any fluctuation in the red curve. As seen, this curve (green) contains almost the same number of maxima and minima as the black curve. The positions of the maxima and minima also coincide with the black curve. That is, the pattern of both curves is almost the same.

The middle and bottom figures contained in Fig.2.1 were plotted for the chains B and C of α subunits. Similar to the top figure for the chain A, these two figures also possess three curves. In the middle figure for the chain B, the curve (red) for the fluctuations of TA2F$_1$ protein is parallel to the residue index axis except around the 400 position in which a peak appears. In the expanded version (green curve), this peak appears to be the largest one among all other maxima. On the same scale of the red curve obtained for the TA2F$_1$ protein, the black curve for the reference protein MF$_1$ contains many maxima and minima, the largest fluctuations occur after the 400 position of the residue index axis. The value of the fluctuations for MF$_1$ protein (black curve) is always higher from amino terminal to the carboxy terminal regions of the chain B (residues 0~500) compared to the fluctuations obtained for the TA2F$_1$ protein except near 300 position where a minima of zero fluctuations appears that coincide with the red curve. The largest variation of the fluctuations in the red and black curves of these two proteins is about 2.5Å^2 near the 500 position in the carboxy terminal region of the chain B. The amino terminal of the B chain for the TA2F$_1$ protein started at 27 and ended at 500 in the carboxy terminal whereas the amino terminal in the same chain for the reference MF$_1$ protein started at 24 and ended at 510 in the carboxy terminal region. Thus the number of amino acids in the B chain for the TA2F$_1$ protein and MF$_1$ protein differs by a factor of 13 and the amino terminal of the B chain of the reference MF$_1$ protein leads by 3 amino acid residues compared to the amino terminal of the B chain of the TA2F$_1$ protein. It seems to be that the maxima and the minima contained in the B chain for both the proteins do not coincide rather they reside in opposite positions, although in the case of chain A, the maxima and minima for both the proteins coincide. However, if the curve for the TA2F$_1$ protein (say green) is moved toward the (amino terminal end) left on the residue axis, the trend of maxima and minima is almost similar to that of the reference protein MF$_1$ up to around the 400 position. After this position, the trend is totally different in the carboxy terminal region which means that the black curve (MF$_1$) although contains several minima but it shows its highest level of fluctuations. In the case of TA2F$_1$ protein, this region is almost flat exhibiting zero fluctuations except a little maximum at around the 435 position.

The red curve (27GLU-500PRO) in the bottom figure obtained for the chain C of α subunit represents the fluctuations of TA2F$_1$ protein. The fluctuation is very little throughout the curve when compared with the black curve (19ALA-510ALA) containing a fluctuation range 0~1.0Å^2 plotted for the MF$_1$ protein on the same scale. In the expanded curve (green), it is clear that some maxima and minima do not coincide but both the curves (black and green) show the similar trend.
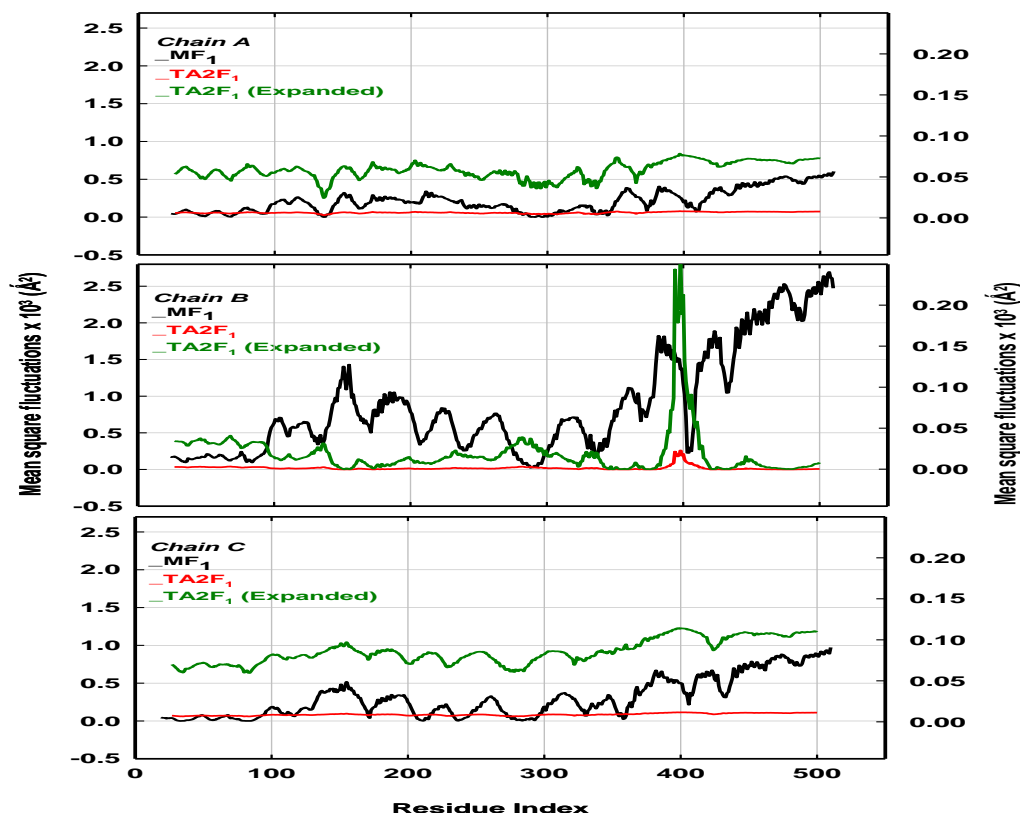
Figure 2.1: The mean-square fluctuations exhibited by different chains of α subunit due to the 1st slowest mode were plotted as a function of residue index. The mean-square fluctuations were multiplied by 1000. The top, middle and bottom curves are for the chains A, B and C respectively. The red curve is for the TA2F$_1$ protein, the black curve is for the MF$_1$ protein and the green curve is an expanded version of the red curve. The fluctuations in right axis are for the green curve. The data were plotted using Sigma Plot.

2.1.2 Analysis of the mean square fluctuations for different chains of the β subunit

The chains (D, E and F) of the β subunit also undergo fluctuations as shown in Fig.2.2. The top figure contains three curves. It appears that the red curve (2ASN-462LEU) obtained by plotting the data for the D chain of TA2F$_1$ protein shows very little but constant fluctuations compared to the black curve (9THR-475GLU) for the MF$_1$ protein. This contains ~ 0.5 Å^2 fluctuations higher than the red curve. The green curve is an expansion of the red curve which shows similar fluctuations as the black curve. A maximum of green curve near 400 residue position seems not to be similar to the black curve because a minimum is observed in the black curve around this residue position.

The middle figure is for the E chain. The red curve (2ASN-462LEU) for the TA2F$_1$ protein exhibits almost the same fluctuations up to 360 residue position compared to the black curve (9THR-474ALA) for the MF$_1$ protein. A peak appears at 383 residue position and then the curve is flat up to the end. From the start to the end, the red curve contains fluctuations in the range of 0~0.75Å^2. The black curve have 0~0.5Å^2 fluctuations and involves some maxima and minima throughout the length of the curve. In the expanded green curve, the maxima and minima contained in the red curve are shown. When the green and red curves are compared, it appears that the fluctuation pattern of the E chain of TA2F$_1$ protein differs from MF$_1$ protein because the positions of the maxima and minima are not the same.

The three curves in the bottom figure were plotted for the F chain of the β subunit. The black curve (10THR-474ALA) for the MF$_1$ protein have fluctuations in the range of 0~0.46Å^2. The red curve (2ASN-462LEU) for the TA2F$_1$ protein is parallel to the residue axis. The green curve shows the fluctuations contained in the red curve. Comparing the black and green curves, it is evident that the fluctuation pattern in the F chain of both the proteins is similar.
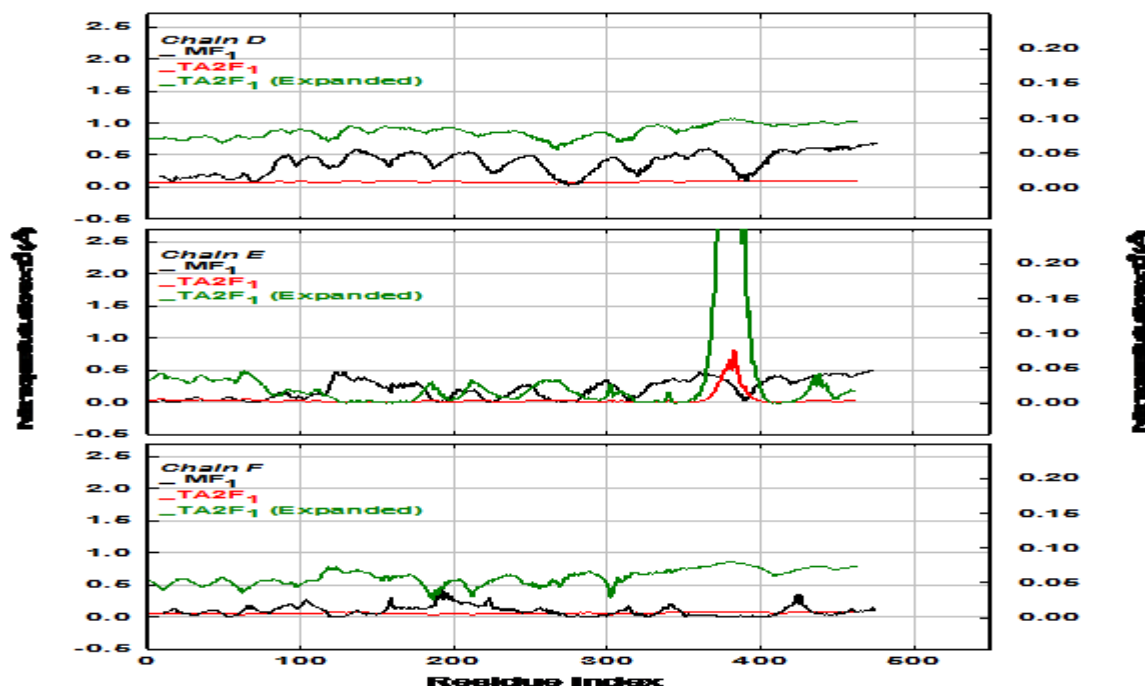

Figure 2.2: The mean-square fluctuations exhibited by different chains of β subunit due to the 1st slowest mode were plotted as a function of residue index. The mean-square fluctuations were multiplied by 1000. The top, middle and bottom curves are for the chains D, E and F respectively. The red curve is for the TA2F$_1$ protein, the black curve is for the MF$_1$ protein and the green curve is an expanded version of the red curve. The fluctuations in right axis are for the green curve. The data were plotted using Sigma Plot.

2.1.3 Analysis of the mean square fluctuations for the γ subunit

The γ subunit of both the TA2F$_1$ and MF$_1$ proteins contain only one chain which is denoted by the letter G. The fluctuations exhibited by the G chain are displayed in Fig.2.3. The red curve (3GLY-266ALA) obtained for the TA2F$_1$ protein shows fluctuations in the range of 0~2.7Å^2 which is much higher than the fluctuations exhibited by the six chains (A, B, C and D, E, F) of α and β subunits (Figs.2.1-2.2). The black curve (1ALA-272LEU) is obtained for the G chain of MF$_1$ protein and the cyan curve is an expansion of the black curve. In the cyan curve, there is a very little fluctuation in the amino terminal around the 1-20 residue positions and at the end of the carboxy terminal of the G chain which is insignificant when compared with the fluctuations in the red curve. It is clear from the figure that the motion of the amino acid residues of G chain of TA2F$_1$ protein is far different from that of MF$_1$ protein.
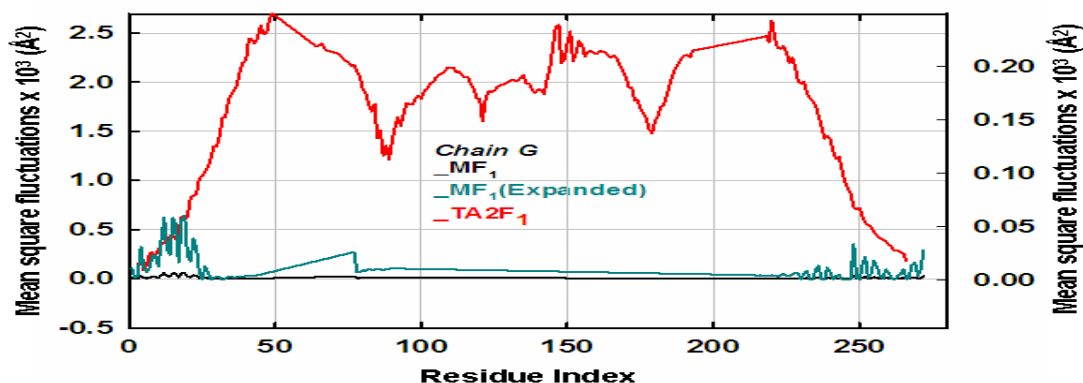
Figure 2.3: The mean-square fluctuations as a function of residue index, driven by the 1st slowest mode for γ subunit. The mean-square fluctuations were multiplied by 1000. The red curve is for the TA2F$_1$ protein, the black curve is for the MF$_1$ protein and the cyan curve is an expanded version of the black curve. The fluctuations in right axis are for the green curve. The data were plotted using Sigma Plot.

2.1.4 Analysis of the mean square fluctuations for the ε subunit

The structure of TA2F$_1$ protein contains the subunit ε having one chain which is expressed by the letter H.The fluctuation of this subunit is plotted in red color as shown in Fig.2.4 and it has a fluctuation of ~3.7Å^2 which is the highest level of fluctuation among all the chains (A, B, C, D, E, F and G) of other subunits ($\alpha_3\beta_3$).
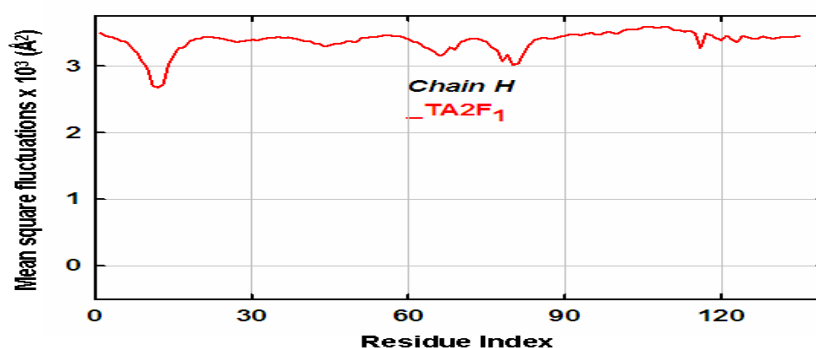

Figure 2.4: The mean-square fluctuations of the ε subunit (H chain) of TA2F$_1$ protein were plotted as a function of residue index. The protein motion was driven by the 1st slowest mode. The mean-square fluctuations were multiplied by 1000. The data were plotted using Sigma Plot.

2.2 Analysis of B factors

The GNM theory was used to calculate the B factors or Debye-Waller factors or temperature factors of individual residues of different chains of MF$_1$ and TA2F$_1$ proteins. The calculated B factors were compared with the experimentally measured B factors as reported in the PDB files (PDB ID:1bmf for MF$_1$ and PDB ID:2QE7 for TA2F$_1$) as shown in Figs.2.5-2.12. The B-factors are related to the mean square fluctuations of individual residues according to the equation 4.10 in chapter IV. The calculated correlation coefficient for MF$_1$ was 0.71 (71%) and for TA2F$_1$ was 0.64 (64%). The correlation coefficient for ribonuclease T1 was calculated as ~0.7 (70%) by GNM theory in a previous work [30].

In general, the overall agreement of calculated B factors with the experimental values for the chains A, B, C, D, E and F of both the MF$_1$ and TA2F$_1$ proteins is very good although there are some deviations in calculated maxima and minima from experimental ones for individual residues (Figs.2.5-2.10). Some examples are given here. In the case of chain A (blue and pink curves) of MF$_1$ protein, there are some deviation of theoretical maxima around 222ASP, 315ALA, 368GLY and 510ALA and minima around 323ALA and 403PHE from experimental ones. For the same chain (dark cyan and dark yellow curves) of TA2F$_1$ protein, significantly deviated maxima and minima are 57ASN, 69ASP, 310GLY and 456PHE, 482PRO, 492GLU respectively. For other chains (B, C, D, E and F) of MF$_1$ protein, it is seen from the top figures that maxima: 222ASP, 314ASP, 360GLY and 506ALA (chain B), 20ASP, 58GLY, 222ASP, 314ASP and 510ALA (chain C), 41ARG, 208LEU, 361ASN and 450ASP (chain D), 40GLY, 209LYS and 473LEU (chain E) and 40GLY, 209LYS and 360PRO (chain F) deviate significantly. Minima: 35GLY and 341ASN (chain B), 274GLN and 350ILE (chain C), 157GLY (chain D), 186VAL and 459MET (chain E) and 321ALA (chain F).For other chains (B, C, D, E and F) of TA2F$_1$ protein, it is seen in the bottom figures that the deviated maxima are: 61GLY, 328ALA and 448ASP (chain B), 57ASN, 214GLN and 466ALA (chain C), 105ALA and 267ALA (chain D), 261LEU, 330GLU and 440ASN (chain E), 72ASP, 261LEU and 353ALA

(chain F). Some deviated minima are: 457GLU, 474HIS and 492GLU (chain B), 163GLN and 233PRO (chain C), 24PRO, 409VAL and 450GLY (chain D), 208PHY (chain E), 450GLY (chain F).
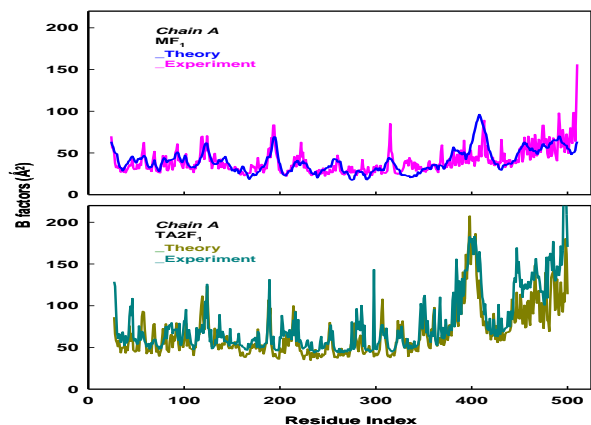

Figure 2.5: Comparison of theoretical and experimental B-factors for the chain B of α subunit of MF_1 (upper figure) and $TA2F_1$ (lower figure) proteins. B-factors are plotted as a function of residue index. The theoretical data were obtained from calculation using oGNM server [26] and experimental data were reported in the PDB files (ID:1bmf for MF_1 and ID:2QE7for $TA2F_1$). Sigma Plot was used to produce the curves.
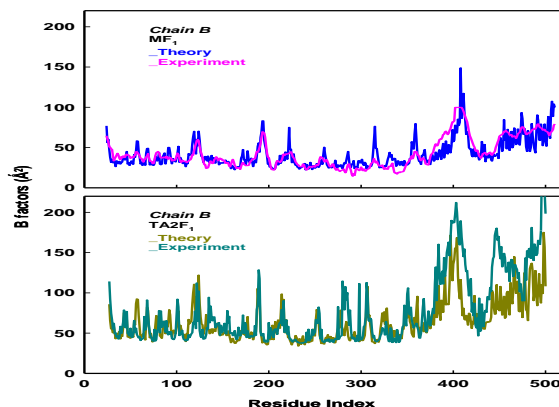
Figure 2.6: Comparison of theoretical and experimental B-factors for the chain B of α subunit of MF_1 (upper figure) and $TA2F_1$(lower figure) proteins. B-factors are plotted as a functionof residue index. The theoretical data were obtained from calculation using oGNM server [26] and experimental data were reported in the PDB files (ID:1bmf for MF_1 and ID:2QE7 for $TA2F_1$). Sigma Plot was used to produce the curves.
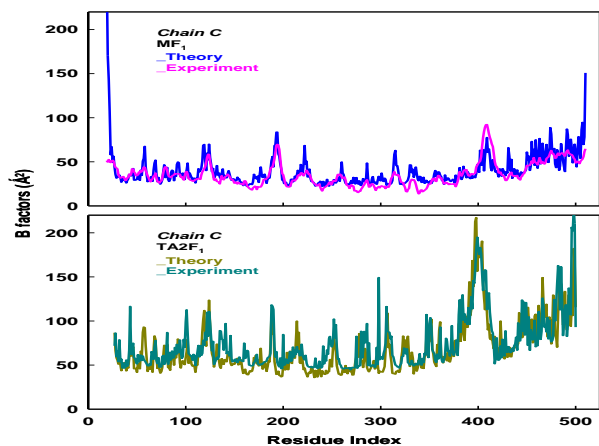

Fig.2.7.Comparison of theoretical and experimental B-factors for the chain C of α subunit of MF_1 (upper figure) and $TA2F_1$ (lower figure) proteins. B-factors are plotted as a function of residue index. The theoretical data were obtained from calculation using oGNM server [26] and experimental data were reported in the PDB files (ID:1bmf for MF_1 and ID:2QE7 for $TA2F_1$). Sigma Plot was used to produce the curves.
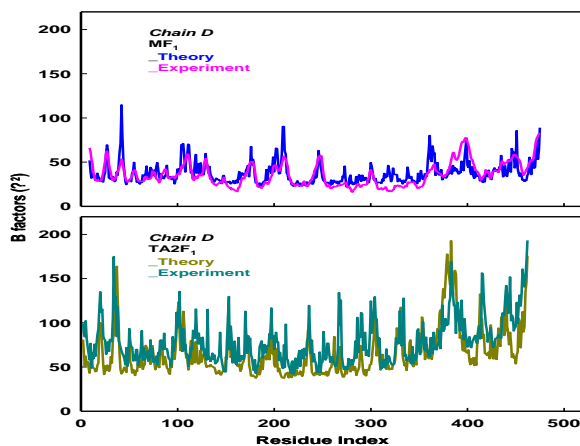
Fig.2.8.Comparison of theoretical and experimental B-factors for the chain D of β subunit of MF_1 (upper figure) and $TA2F_1$ (lower figure) proteins B-factors are plotted as a function of residue index. The theoretical data were obtained from calculation using oGNM server [26] and experimental data were reported in the PDB files (ID:1bmf for MF_1 and ID:2QE7 for $TA2F_1$). Sigma Plot was used to produce the curves.
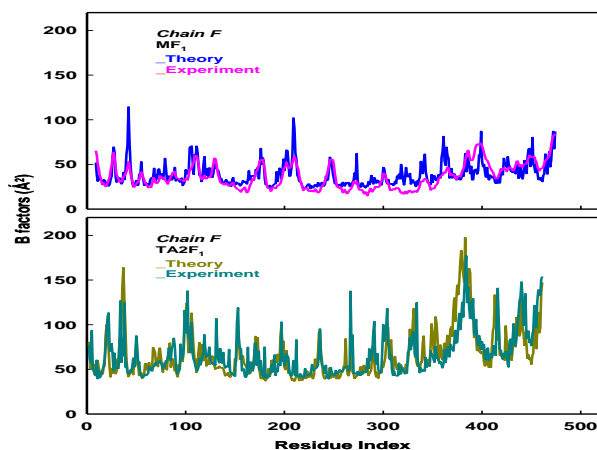
Fig.2.9.Comparison of theoretical and experimental B-factors for the chain E of β subunit of MF$_1$ (upper figure) and TA2F$_1$ (lower figure) proteins. B-factors are plotted as a function of residue index. The theoretical data were obtained from calculation using oGNM server [26] and experimental data were reported in the PDB files (ID:1bmf for MF$_1$ and ID:2QE7 for TA2F$_1$). Sigma Plot was used to produce the curves.
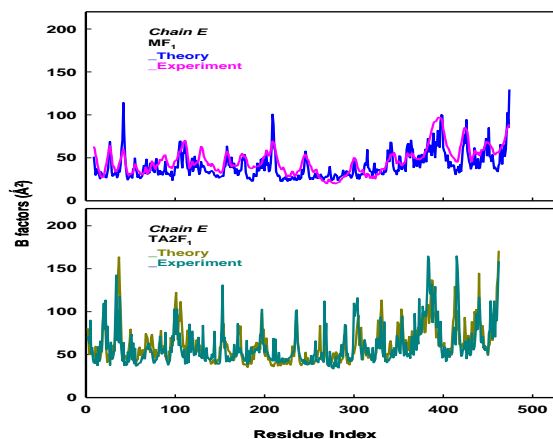
Fig.2.10.Comparison of theoretical and experimental B-factors for the chain E of β subunit of MF$_1$ (upper figure) and TA2F$_1$ (lower figure) proteins. B-factors are plotted as a function of residue index. The theoretical data were obtained from calculation using oGNM server [26] and experimental data were reported in the PDB files (ID:1bmf for MF$_1$ and ID:2QE7 forTA2F$_1$). Sigma Plot was used to produce the curves.
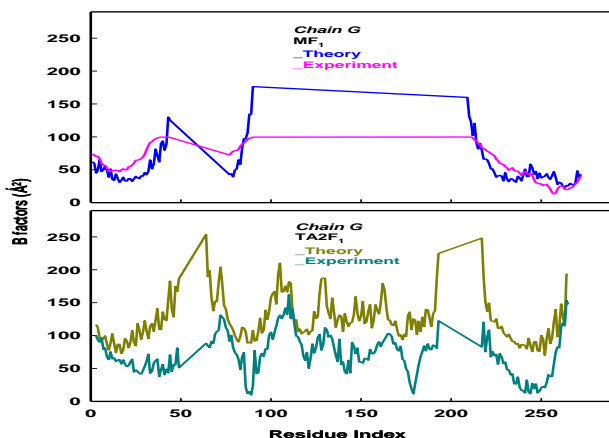

Fig.2.11.Comparison of theoretical and experimental B-factors for the chain D of β subunit of MF$_1$ (upper figure) and TA2F$_1$ (lower figure) proteins. B-factors are plotted as a function of residue index. The theoretical data were obtained from calculation using oGNM server [26] and experimental data were reported in the PDB files (ID:1bmf for MF$_1$ and ID:2QE7 for TA2F$_1$). Sigma Plot was used to produce the curves.

Fig.2.12.Comparison of theoretical and experimental B-factors for the chain D of β subunit of MF$_1$ (upper figure) and TA2F$_1$ (lower figure) proteins. B-factors are plotted as a function of residue index. The theoretical data were obtained from calculation using oGNM server [26] and experimental data were reported in the PDB files (ID:1bmf for MF$_1$ and ID:2QE7 for TA2F$_1$). Sigma Plot was used to produce the curves.
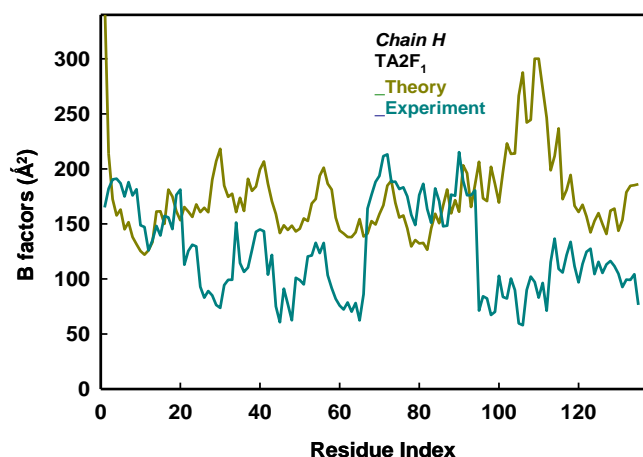
The pattern of B factor curves in Fig.2.11 obtained for the G chain (γ subunit) of both the MF$_1$ and TA2F$_1$ proteins are far different from other curves in Figs.2.5-2.10. In the pink curve drawn from experimental data reveals a minimum around 8ARG-21LYS, a maximum around 35GLU-44TYR and a minimum around 77LEU-81ILE. The blue curve obtained by plotting the data obtained by GNM calculation is almost similar to that of the experimental curve but there are very high deviations in maxima, minima or flat regions. Although the experimental and theoretical curve pattern for TA2F$_1$ protein is similar but the calculated B factors are much higher than those of experimental ones. Remarkably deviated maxima and minima in the G (γ subunit) chain are: 65HIS, 130GLY, 162ASP and88ALA, 178PRO, 245MET respectively. Similar to the chain G, the theoretical B factors for chain H (ε subunit) are very high compared to the experimental values. Some highly deviated maxima and minima in this chain are: 30GLY, 110THR and 106LYS and 64SER respectively.

III.   DISCUSSION

In the structure of mitochondrial F$_1$-ATPase (MF$_1$) determined in the presence of nucleotide and magnesium, three α and three β subunits are alternately arranged to form a α$_3$β$_3$ cylinder in which the subunits are folded identically [7]. An antiparallel α-helical coiled coil of amino- (N-) and carboxy- (C-) termini of the γsubunit forms the axle of rotation that deeply penetrates the α$_3$β$_3$ cylinder of F$_1$-ATPase. This axle is held at the top orifice and the bottom of the α$_3$β$_3$ cylinder as shown in Fig.2.13 for MF$_1$ protein. The bottom part is made of a hydrophobic ''sleeve'' that acts as a bearing [7]. The structure of α$_3$β$_3$ cylinder of F$_1$-ATPase from thermophilic*Bacillus* PS3 [41] was found to be similar to that of MF$_1$. The αand β subunits of TA2F$_1$ [7] forming the α$_3$β$_3$ cylinder have a similar overall fold as MF$_1$ and TF$_1$ structures [7, 41]. It is thus clear that the structures of α$_3$β$_3$ subunits of the three F$_1$-ATPases as mentioned above are similar. Each α or β subunit is composed of an N-terminal six stranded β-barrel, a central domain containing nucleotide binding sites and a C-terminal domain [7].

The fluctuations exhibited by different chains of α (A, B, C) and β (D, E, F) subunits of TA2F$_1$ protein is very low compared to the fluctuations of different chains of the MF$_1$ protein as observed in red and black curves in Figs.2.1-2.2. The fluctuations hidden in the red curves are revealed in the expanded green curves. In the fluctuation curves of B and E chains of TA2F$_1$ protein, very high peaks appeared near 400 residue position. Considering these points, it seems to be that the structures of α and β subunits of TA2F$_1$ protein are very rigid compared to the MF$_1$ protein. The differences in fluctuations of these two proteins may be due to their structural differences. The structure of TA2F$_1$ protein was determined in the absence

of nucleotides and magnesium so that the α and β subunits could not bind any nucleotide and they remained in open conformations. The structure of MF_1 protein was determined in the presence of nucleotides and magnesium and consequently the α and β subunits remained in closed conformation by binding Mg-nucleotide except one β subunit which remained empty and open.
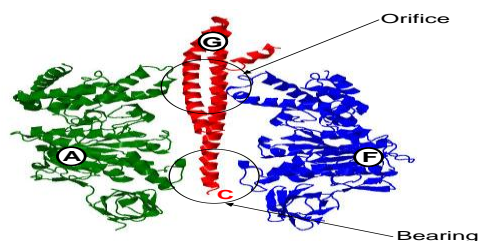

Fig.2.13. The chain A of α subunit and the chain F of β subunit together with γ subunit (chain G) of MF_1 [26] are shown. The chains are represented by letters in the circles. The carboxy (C) terminal part of the γ subunit deeply penetrates into the molecular bearing made of the bottom part of the $α_3β_3$ subunits. The upper and lower circles represent the entrance (orifice) and bearing regions of the $α_3β_3$ cylinder. The figure is made using Jmol software [34].

In the case of MF_1, the amino terminal regions through the A-F chains of αβ subunits around 24-95 residues is almost rigid (Figs.2.1-2.2). The regions 100-400 residues in each fluctuation curve are more flexible compared to the amino terminal region. The regions around 400~500 residues containing carboxyl terminal regions show the highest flexibility compared to the other two regions. The curve for the chain B exhibits the highest mobility and the region around 400~500 residues in the curve for the chain F exhibits less flexibility. The α subunits are more flexible compared to the β subunits.

In some mutational studies [31, 32], the residues 155LYS and 182ARG in the β subunit of the *Escherichia coli*F_1-ATPase (EF_1) corresponding to the residues 162LYS and 189ARG in the β subunit of MF_1 were found important for generation and stabilization of the catalytic transition state. The residue β-182ARG plays role in the required conformational change at the catalytic sites for multisite catalysis [34, 35].The residues 162LYS and 189ARG in the β subunit of MF_1 are located within the minima in the black curves obtained for chains D, E and F in Fig.2.2. These minima have some fluctuations i.e. flexibility. The residues in the green curve for D, E and F chains of $TA2F_1$ protein corresponding to the residues 162LYS and 189ARG in MF_1 also lie in minima but with some fluctuations/flexibility region. These similarities indicate that some flexibility/conformational changes are necessary for catalysis in F_1-ATPases irrespective of their origins. The residue 188GLU in the β subunit of MF_1 [19] lies in a position to activate the water molecule that enhances a nucleophilic attack on the terminal phosphate. In the case of F_1-ATPase from *Bacillus* PS3 (TF_1), mutations of βE190D and βE190DCax residues (corresponding to β188GLU in MF_1) severely impaired catalysis [40]. The position of the residue β188GLU in the black curves for D, E and F chains of MF_1 is within minima with some fluctuations. The position of the residue in the green curves for $TA2F_1$ protein corresponding to β188GLU in MF_1 remains in the regions having some fluctuations. Thus it is evident that some sorts of molecular motion or change in conformation occur in the relevant sites of subunits of F_1-ATPase for nucleotide hydrolysis.

Considering the overall fluctuations in the Figs.2.1-2.2, it is clear that α and β subunits of $α_3β_3$ cylinder in $TA2F_1$ protein are highly rigid compared to the flexible $α_3β_3$ cylinder of MF_1 protein. The γ (G) subunit of MF_1 protein almost fully penetrates into the cavity of $α_3β_3$ cylinder up to the molecular bearing (Fig.2.13) and its motion may be highly restricted by the $α_3β_3$ cylinder so that it shows very little fluctuations (Fig.2.3) compared to the $α_3β_3$ subunits. In the case of $TA2F_1$ protein, only a part of the γ (G) subunit of $TA2F_1$ protein enter through the orifice and it remains in the upper portion of the cavity of the $α_3β_3$ cylinder (Fig.2.14). Rest of the cavity remains empty. Moreover, the protruding portion of the γ (G) subunit is attached to the ε (H) subunit. The γ (G) subunit of $TA2F_1$ protein thus feels very low level of restrictions from $α_3β_3$ cylinder around it and this may be an important reason for very high fluctuations of the γ (G) subunit. The ε (H) subunit remains completely outside the cavity of $α_3β_3$ cylinder so that it gets no resistance from $α_3β_3$ cylinder and it can move more freely during molecular motion. This may be one reason why ε (H) subunit shows exceptionally high fluctuations compared to the all other subunits of $TA2F_1$ protein and that of MF_1 protein.

Fig.2.14. The chain A of α subunit and the chain F of β subunit together with γ subunit (chain G) and ε subunit (chain H) of $TA2F_1$ [24] are shown. The chains are represented by letters in the circles. The carboxy (C) terminal part of the γ subunit remains in the cavity region of the $α_3β_3$ cylinder. The upper and lower circles represent the entrance (orifice) and bearing regions of the $α_3β_3$ cylinder. The figure is made using Jmol software [34].

The more localized atoms generate relatively low B factor values and they usually take part to form secondary structures, neighboring disulfide bridges, etc. The chain termini exhibit fluctuation peaks in several proteins due to their inherent enhanced flexibility. The surface atoms generally exhibit larger fluctuations, compared to core atoms. In the case of a multimer protein consisting of several subunits, theoretical curves fall above the experimental curves at the regions of subunit-subunit interfaces, due to the absence of interactions between neighboring subunits [9].

The γ subunit of F_1-ATPase rotates inside the cylinder made of $\alpha_3\beta_3$ subunits and rotation was first observed in TF_1 [26] using various probes such as actin filament [26], a spherical bead or a bead duplex [30, 31] and so on. Subsequently, rotation was also observed in F_1-ATPases from *Escherichia coli (EF$_1$)*and from chloroplast (CF_1) [32, 33]. Rotational experiment on $TA2F_1$ protein has not yet been reported. According to a push pull mechanism [51], when a β subunit binds a nucleotide its upper portion bends toward and apparently pushes the axle at the top and when the β subunit releases the nucleotide it pulls the axle. The push pull actions generate rotation of the slightly curved axle in a conical fashion. This mechanism requires that the axle to be rigid and its tip held in the bearing acts as a pivot.

In MF_1, the sleeve/bearing consists of the six proline-rich loops (α287-294 and β274-281) of the $\alpha_3\beta_3$ cylinder. The residues 287-294 remain in the minima regions of black curves obtained for chains A, B and C of α subunit as seen in Fig.2.1. Comparing the green curves of chains A, B and C of α subunit of $TA2F_1$ protein in the same figure it is observed that the corresponding residues for $TA2F_1$ protein also lie in the minima of green curves. Similarly, the residues 274-281 in the β subunit of MF_1 remain in the minima of black curves and the corresponding residues in $TA2F_1$ protein lie almost in the minima of green curves as shown in Fig.2.2 plotted for the chains D, E and F. Thus from the study of molecular motion it appears that the residues forming the molecular bearing of both MF_1 and $TA2F_1$ proteins remain in the minima which implies that bearing is quite rigid that helps to hold the carboxy (C) terminal tip of the axle giving an overall structural stability of F_1-ATPases. However, an axle-less mutant of F_1-ATPase from thermophilic bacterium (TF_1) could rotate for >100 revolutions [29] which implies that neither a rigid axle nor a fixed support at the bottom is necessary for rotation.

The highly conserved and charged DELSEED sequence (β394ASP-400ASP) of β subunit that interacts with the γsubunit can be replaced with alanine without affecting rotational torque [28] giving an impression that the interactions between the protruding portion of γ subunit and the top surface of β subunit is less important. As seen in Fig.2.2, the residues 394ASP-400ASP of the β subunit of MF_1 represented by black lines for the chains D, E and F lie within minima. The corresponding residues for the $TA2F_1$ protein also remain in the minima regions of green curves. This suggests that these residues are restricted or least mobile.

IV.   CONCLUSION

F1-ATPase is a motor protein that generates rotation during ATP hydrolysis. A motor protein is a machine that converts chemical energy into mechanical energy and forms the basis of a specific cellular function. The proteins undergo motion and exhibit different ranges of flexibility during their functions. The molecular motions of two proteins of different sources namely, F1-ATPase from bovine mitochondria (MF1) and F1-ATPase from thermoalkaliphilic bacteria (TA2F1) were studied in this thesis using Gaussian Network Model (GNM). The data obtained by calculation using GNM server [26] for MF1 and TA2F1 proteins for the 1st slowest mode of vibration were compared and analyzed. A force constant $\gamma = 0.054$ kcal/(mole·Å^2) was obtained for TA2F1 protein and $\gamma = 0.08$ kcal/(mole·Å^2) [27] was for MF1 protein. The variation in the force constant is =0.026 kcal/(mole·Å^2) which implies that the inter residue contact potential is stronger in MF1 than in TA2F1. The degree of fluctuations of the different chains of α subunit (A, B and C) and β subunit (D, E and F) of the MF1 and TA2F1 proteins varies from 0~ 2.5$\text{Å}2$ (Figs.2.1-2.2). The largest fluctuations ~ 2.5 Å^2 in MF1 occurs in the carboxy terminal region of the chain B. This suggests that $\alpha_3\beta_3$ the cylinder of MF1 is flexible whereas the $\alpha_3\beta_3$cylinder of TA2F1 is highly rigid. The γ subunit of TA2F1 remains in the cavity of the $\alpha_3\beta_3$cylinder near the orifice (Fig.2.14) and it gets less resistance during motion compared to the γ (G) subunit of MF1 (Fig.2.13). This may cause a fluctuation ~ 2.5Å^2 higher in TA2F1 than in MF1. The fluctuation of ε (H) subunit of TA2F1 tagged with the protruding portion of the γ (G) subunit is the highest ~3.5 Å^2 compared to all other subunits of both the proteins. The overall agreement between theoretical and experimental values of B factors is good which is supported by the correlation coefficient. The calculated correlation coefficient for MF1 was 0.71 (71%) [27] and for TA2F1 was 0.64 (64%). The residue β-188GLU in MF1 responsible for ATP hydrolysis andthe corresponding residue in TA2F1 exhibit flexibility. The hydrophobic 'sleeve'44 forming the molecular bearing was found to be located in the minima for both the proteins.

REFERENCES

[1]   Mitchell, P. Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature.*191, 144–148(1961).
[2]   Elston, T., H. Wang and G. Oster. Energy transduction in ATP synthase.*Nature.*391, 510-513(1998).
[3]   Boyer, P.D.The binding change mechanism for ATP synthase—some probabilities and possibilities.*Biochim.Biophys.Acta.*1140, 215–250 (1993).
[4]   Kinosita, K. Jr, R. Yasuda, H. Noji& K. Adachi.A rotary molecular motor that can work at near 100% efficiency.*Philos. Trans. R. Soc. Lond. B. Biol. Sci.*355, 473–489(2000).
[5]   Yoshida, M., E. Muneyuki& T. Hisabori. ATP synthase—a marvellous rotary engine of the cell. *Nat. Rev. Mol. Cell Biol.*2, 669–677(2001).

[6]   http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B94RW-4VB4W2R-17&_user=133418&_coverDate=11%2F15%2F2008&_rdoc=1&_fmt=high&_orig=search&_sort=d&_docanchor=&view=c&_acct=C000010718&_version=1&_urlVersion=0&_userid=133418&md5=37ee4b867a04c34cbf02a0558e75661b"   \l "bbib7#bbib7Kinosita, K. Jr, K. Adachi & H. Itoh. Rotation of F_1-ATPase: how an ATP-driven molecular machine may work. *Annu. Rev. Biophys. Biomol.Struct.*33, 245–268(2004).

[7]   http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B94RW-4VB4W2R-17&_user=133418&_coverDate=11%2F15%2F2008&_rdoc=1&_fmt=high&_orig=search&_sort=d&_docanchor=&view=c&_acct=C000010718&_version=1&_urlVersion=0&_userid=133418&md5=37ee4b867a04c34cbf02a0558e75661b" \l "bbib8#bbib8 PDB ID:1bmf. Abrahams, J.P., A.G.W. Leslie, R. Lutter and J.E. Walker.Structure at 2.8 Å resolution of F_1-ATPase from bovine heart mitochondria.*Nature.*370, 621–628(1994).

[8]   Yagi H., N. Kajiwara, H. Tanaka, T. Tsukihara, Y. K. Yamada, M. Yoshida and H. Akutsu. Structures of the thermophilic F1-ATPase εsubunit suggesting ATPregulatedarm motion of its C-terminal domain in F1. *PNAS.*USA. 104, 11233–11238 (2007).

[9]   Kato Y., T. Matsui, N. Tanaka, E. Muneyuki, T. Hisabori and MasasukeYoshida. Thermophilic F1-ATPase Is activated without dissociation of an endogenous inhibitor, ε subunit. *J. of Biol. Chem.* 272, 24906–24912(1997).

[10]   http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B94RW-4VB4W2R-17&_user=133418&_coverDate=11%2F15%2F2008&_rdoc=1&_fmt=high&_orig=search&_sort=d&_docanchor=&view=c&_acct=C000010718&_version=1&_urlVersion=0&_userid=133418&md5=37ee4b867a04c34cbf02a0558e75661b"   \l "bbib12#bbib12Noji, H., R. Yasuda, M. Yoshida & K. Kinosita Jr. Direct observation of the rotation of F_1-ATPase.*Nature.*386, 299–302(1997).

[11]   Itoh H., A. Takahashi, K. Adachi, H. Noji, R. Yasuda, M. Yoshida and K. Kinosita Jr. Mechanically driven ATP synthesis by F1-ATPase. *Nature.*427, 465- 468 (2004).

[12]   Liu X. and H. A. Karimi.High-throughput modeling and analysis of protein structural dynamics.*Briefing.in Bioinformat*.8. 432-445 (2007).

[13]   Bahar I, A.R. Atilgan and B. Erman. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential.*Fold Des*.2, 173–181(1997).

[14]   Haliloglu T, I. Bahar and B. Brman. Gaussian dynamics of folded proteins.*Phys Rev Lett* .79, 3090–3093(1997).

[15]   Rader A.J., C. Chennubhotla, L.-W.Yang and I. Bahar.The Gaussian Network Model: Theory and Applications. BICH: c472x_c003, 41-64 (2005).

[16]   Tama, F. and Y. H. Sanejouand. Conformational change of proteins arising from normal mode calculations.*Protein Eng*., [14] 1–6(2001).

[17]   Abseher, R. and M. Nilges. Efficient sampling in collective coordinate space.*Proteins*.39, 82–88(2000).

[18]   Keskin O., R. Stewart, Durell, I. Bahar, L. Robert, Jernigan and G. D. Covell.Relating molecular flexibility to function: A case study of tubulin.*Biophys. J.* 83, 663-680 (2002).

[19]   Bahar, I., A. R. Atilgan, M. C. Demirel and B. Erman. Vibrational dynamics of folded proteins: significance of slow and fast motions in relation to function and stability. *Phys. Rev. Lett.*80, 2733–2736(1998.).

[20]   Brooks, B. R., D. Janezic and M. Karplus. Harmonic analysis of large systems.*Methodology. J. Comput. Chem.*16, 1522–1542(1995.).

[21]   Go, N. A theorem on amplitudes of thermal atomic fluctuations in large molecules assuming specific conformations calculated by normal mode analysis.*Biophys.Chem.* 35,105–112(1990).

[22]   Isin, B., P. Doruker and I. Bahar. Functional motions of influenza virus hemagglutinin: a structure-based analytical approach. *Biophys. J.*82, 569–581(2002).

[23]   Kundu S, J.S. Melton, D.C. Sorensen and G.N. Phillips. Dynamics of proteins in crystals: comparison of experiment with simple models. *Biophys.J.*83, 723–732(2002).

[24]   PDB ID:2QE7. Stocker A., S. Keis, J. Vonck, G. M. Cook and P. Dimroth.The structural basis for unidirectional rotation of Thermoalkaliphilic F_1-ATPase.*Structure*, 15, 904–914(2007).

[25]   Berman H.M., J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne (2000) The Protein Data Bank. Nucleic Acids Research, 28: 235-242.

[26]   Yang L.-W., A. J. Rader, L. Xiong, C.J. Jursa, S. C. Chen, H. A. Karimi and I. Bahar. oGNM: online computation of structural dynamics using the Gaussian Network Model. *Nucleic Acids Research*.34, W24–W31(2006).

[27]   Asaduzzaman M. A study on the molecular motion of F1-ATPase using Gaussian Network model. A thesis submitted to the Department of Physics, Shahjalal University of Science and Technology, Sylhet as a partial fulfillment for the Degree of Master of Science. Session:2008-2009.

[28]   Hara, K. Y., H. Noji, D. Bald, R. Yasuda, K. Kinosita Jr. and M. Yoshida. The role of the DELSEED motif of the βsubunit in rotation of F1-ATPase.*J.Biol.Chem.* 275, 14260–14263(2000).

[29]   Keskin, O., R. L. Jernigan and I. Bahar. Proteins with similar architecture exhibit similar large-scale dynamic behavior. *Biophys. J.* 78, 2093–2106 (2000).

[30]   Demirel M. C., A. R. Atilgan, R. Jernigan, B. Erman and I. Bahar. Identification of kinetically hot residues in proteins.*Protein Science.*7, 2522- 2532(1998).

[31]   Noji, H., K. H˙asler, W. Junge, K. Kinosita Jr., M. Yoshida and S. Engelbrecht. Rotation of *Escherichia coli* F1-ATPase.*Biochem.Biophys.Res.Commun.* 260, 597–599(1999).

[32]   Amano T., K. Tozawa, M. Yoshida and H. Murakami. Spatial precision of a catalytic carboxylate of F1-ATPase γ subunit probed by introducing different carboxylate-containing side chains.*FEBS Letters.*348, 93-98(1994).

[33]   Yasuda, R., H. Noji, K. KinositaJr& M. Yoshida. F1-ATPase is a highly efficient molecular motor that rotates with discrete 120° steps. *Cell.*93, 1117–1124(1998).

[34]   Jmol: an open-source Java viewer for chemical structures in 3D.

[35]   Furuike, S., M.D. Hossain, Y. Maki, K. Adachi, T. Suzuki, A. Kohori, H. Itoh, M. Yoshida & K. Kinosita Jr. Axle-less F1-ATPase rotates in the correct direction. *Science.319*, 955–958(2008).